ELSEVIER

Linear Algebra and its Applications 308 (2000) 77–84

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

Matrix Young inequalities for the Hilbert–Schmidt norm

Omar Hirzallah, Fuad Kittaneh*

Department of Mathematics, University of Jordan, Amman, Jordan

Received 12 September 1999; accepted 10 December 1999

Submitted by R. Bhatia

Abstract

Let A, B, and X be $n \times n$ complex matrices such that A and B are positive semidefinite. If $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$, it is shown that $\|\frac{1}{p}A^pX + \frac{1}{q}XB^q\|_2^2 \geqslant \frac{1}{r^2}\|A^pX - XB^q\|_2^2 + \|AXB\|_2^2$, where $r = \max(p, q)$ and $\|\cdot\|_2$ is the Hilbert–Schmidt norm. Generalizations and applications of this inequality are also considered. © 2000 Elsevier Science Inc. All rights reserved.

AMS classification: 15A45; 15A60; 47B10

Keywords: Hilbert–Schmidt norm; Matrix Young inequality; Positive semidefinite matrix

Let $M_n(\mathbb{C})$ denote the space of all $n \times n$ complex matrices. If $A = [a_{ij}] \in M_n(\mathbb{C})$, then the Hilbert–Schmidt (or Frobenius) norm of A is given by

$$\|A\|_2 = \left(\sum_{i,j=1}^{n} |a_{ij}|^2\right)^{1/2}. \qquad (1)$$

It is known that this norm is unitarily invariant, i.e.,

$$\|UAV\|_2 = \|A\|_2 \qquad (2)$$

for all unitary matrices $U, V \in M_n(\mathbb{C})$ (see [3, p. 7] or [7, pp. 291, 292]). Using this, together with the singular value decomposition of A, we have

* Corresponding author.
E-mail address: fkitt@ju.edu.jo (F. Kittaneh).

PII: S 0 0 2 4 - 3 7 9 5 (9 9) 0 0 2 7 0 - 0

$$\|A\|_2 = \left(\sum_{j=1}^{n} s_j^2(A) \right)^{1/2}, \tag{3}$$

where $s_1(A) \geqslant s_2(A) \geqslant \cdots \geqslant s_n(A)$ are the singular values of A, i.e., the eigenvalues of $|A| = (A^*A)^{1/2}$, arranged in decreasing order and repeated according to multiplicity.

The classical Young inequality for nonnegative real numbers says that if $a, b \geqslant 0$ and $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$, then

$$\frac{a^p}{p} + \frac{b^q}{q} \geqslant ab \tag{4}$$

with equality if and only if $a^p = b^q$. Several matrix versions of the Young inequality (4) have been recently established. Kosaki [12] and Bhatia and Parthasarathy [6] pointed out that if $A, B, X \in M_n(\mathbb{C})$ such that A and B are positive semidefinite and if $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$, then

$$\left\| \frac{1}{p} A^p X + \frac{1}{q} X B^q \right\|_2 \geqslant \|AXB\|_2 . \tag{5}$$

Moreover, it has been shown earlier by Ando [2] that if $A, B \in M_n(\mathbb{C})$ are positive semidefinite and if $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$, then

$$s_j \left(\frac{1}{p} A^p + \frac{1}{q} B^q \right) \geqslant s_j \left(AB \right) \tag{6}$$

for $j = 1, 2, \ldots, n$. Equivalently,

$$U \left(\frac{1}{p} A^p + \frac{1}{q} B^q \right) U^* \geqslant |AB| \tag{7}$$

for some unitary matrix U depending on A and B.

Note that inequalities (6) and (7) are much stronger than inequality (5) when $X = I$ (the identity matrix). The special case $p = q = 2$ of inequalities (6) and (7) has been obtained earlier by Bhatia and Kittaneh [4]. It should be mentioned here that for $p = q = 2$, inequality (5), which is a matrix arithmetic–geometric mean inequality, is valid for all unitarily invariant norms. For comprehensive discussions of matrix arithmetic–geometric mean and related inequalities, we refer to [3,5,8,10,11,12] and the references therein.

The main purpose of this paper is to give a refinement of inequality (5). Our refined version of (5) enables us to investigate the equality conditions of inequalities (5)–(7). To achieve our goal we need the following refinement of the classical Young inequality (4).

Lemma 1. *If $a, b \geqslant 0$ and $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$, then*

$$\left(\frac{a^p}{p} + \frac{b^q}{q} \right)^2 \geqslant \frac{1}{r^2} \left(a^p - b^q \right)^2 + a^2 b^2, \tag{8}$$

where $r = \max(p, q)$.

Proof. First observe that for $p = q = 2$, inequality (8) degenerates to an equality. If $q > p$, then $q > 2$, and it follows by direct computations that

$$\left(\frac{a^p}{p} + \frac{b^q}{q}\right)^2 - \frac{1}{q^2}\left(a^p - b^q\right)^2 = a^p\left(\left(1 - \frac{2}{q}\right)a^p + \frac{2}{q}b^q\right).$$

Now using inequality (4), it can be easily seen that

$$\left(1 - \frac{2}{q}\right)a^p + \frac{2}{q}b^q \geqslant a^{p(1-2/q)}b^{q(2/q)} = a^{(q-p)/q}b^2.$$

Consequently,

$$\left(\frac{a^p}{p} + \frac{b^q}{q}\right)^2 - \frac{1}{q^2}\left(a^p - b^q\right)^2 \geqslant a^p a^{(q-p)/q}b^2 = a^2b^2$$

and so

$$\left(\frac{a^p}{p} + \frac{b^q}{q}\right)^2 \geqslant \frac{1}{q^2}\left(a^p - b^q\right)^2 + a^2b^2.$$

By a similar argument, it can be seen that if $p > q$, then

$$\left(\frac{a^p}{p} + \frac{b^q}{q}\right)^2 \geqslant \frac{1}{p^2}\left(a^p - b^q\right)^2 + a^2b^2.$$

Hence,

$$\left(\frac{a^p}{p} + \frac{b^q}{q}\right)^2 \geqslant \frac{1}{r^2}\left(a^p - b^q\right)^2 + a^2b^2$$

as required. □

Our refined matrix Young inequality can be stated as follows.

Theorem 1. *Let $A, B, X \in M_n(\mathbb{C})$ such that A and B are positive semidefinite. If $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$, then*

$$\left\|\frac{1}{p}A^pX + \frac{1}{q}XB^q\right\|_2^2 \geqslant \frac{1}{r^2}\left\|A^pX - XB^q\right\|_2^2 + \|AXB\|_2^2, \tag{9}$$

where $r = \max(p, q)$.

Proof. Since every positive semidefinite matrix is unitarily diagonalizable, it follows that there are unitary matrices $U, V \in M_n(\mathbb{C})$ such that $A = U\Lambda U^*$ and $B = VMV^*$, where $\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, $M = \operatorname{diag}(\mu_1, \mu_2, \ldots, \mu_n)$, and all λ_i, μ_i are nonnegative. If $Y = U^*XV = [y_{ij}]$, then

$$\begin{aligned}\frac{1}{p}A^pX + \frac{1}{q}XB^q &= U\left(\frac{1}{p}\Lambda^pY + \frac{1}{q}YM^q\right)V^* \\ &= U\left[\left(\frac{\lambda_i^p}{p} + \frac{\mu_j^q}{q}\right)y_{ij}\right]V^*,\end{aligned}$$

$$A^pX - XB^q = U\left(\Lambda^pY - YM^q\right)V^* = U\left[\left(\lambda_i^p - \mu_j^q\right)y_{ij}\right]V^*$$

and

$$AXB = U(\Lambda YM)\,V^* = U\left[\lambda_i\mu_j y_{ij}\right]\,V^*.$$

Now using (1), (2), and inequality (8) applied to the nonnegative numbers λ_i, μ_j for $i, j = 1, 2, \ldots, n$, we obtain

$$\begin{aligned}\left\|\frac{1}{p}A^pX + \frac{1}{q}XB^q\right\|_2^2 &= \sum_{i,j=1}^{n}\left(\frac{\lambda_i^p}{p} + \frac{\mu_j^q}{q}\right)^2 \left|y_{ij}\right|^2 \\ &\geqslant \frac{1}{r^2}\sum_{i,j=1}^{n}\left(\lambda_i^p - \mu_j^q\right)^2\left|y_{ij}\right|^2 + \sum_{i,j=1}^{n}\lambda_i^2\,\mu_j^2\left|y_{ij}\right|^2 \\ &= \frac{1}{r^2}\left\|A^pX - XB^q\right\|_2^2 + \|AXB\|_2^2.\end{aligned}$$

This completes the proof of the theorem. □

Inequality (9) makes it possible for us to give necessary and sufficient conditions for equality to satisfy inequalities (5)–(7). This is demonstrated in the following three corollaries.

Corollary 1. *Let $A, B, X \in M_n(\mathbb{C})$ such that A and B are positive semidefinite. If $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$, then*

$$\left\|\frac{1}{p}A^pX + \frac{1}{q}XB^q\right\|_2 = \|AXB\|_2$$

if and only if $A^pX = XB^q$.

Proof. If $A^pX = XB^q$, then by the spectral theorem we have $AX = XB^{q/p}$, and so $\frac{1}{p}A^pX + \frac{1}{q}XB^q = XB^q = XB^{(p+q)/p} = XB^{q/p}B = AXB$.

To prove the "only if" part, assume that $\left\|\frac{1}{p}A^pX + \frac{1}{q}XB^q\right\|_2 = \|AXB\|_2$. Then it follows from inequality (9) that $\|A^pX - XB^q\|_2 = 0$, and so $A^pX = XB^q$, as required. □

Corollary 2. *Let $A, B \in M_n(\mathbb{C})$ be positive semidefinite. If $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$, then*

$$s_j\left(\frac{1}{p}A^p + \frac{1}{q}B^q\right) = s_j(AB) \quad \textit{for } j = 1, 2, \ldots, n$$

if and only if $A^p = B^q$.

Proof. If $A^p = B^q$, then $A = B^{q/p}$, and so $\frac{1}{p}A^p + \frac{1}{q}B^q = B^q = AB$.

To prove the “only if” part, assume that

$$s_j\left(\frac{1}{p}A^p+\frac{1}{q}B^q\right)=s_j\,(AB)\quad \text{for } j=1,2,\ldots,n.$$

Then by (2), we have $\left\|\frac{1}{p}A^p+\frac{1}{q}B^q\right\|_2=\|AB\|_2$, and so it follows from the case $X=I$ of inequality (9) that $A^p=B^q$, as required. □

Using (2) and an argument similar to that used in the proof of Corollary 2, we have the following.

Corollary 3. *Let $A,B\in M_n(\mathbb{C})$ be positive semidefinite. If $p,q>1$ with $\frac{1}{p}+\frac{1}{q}=1$, then*

$$U\left(\frac{1}{p}A^p+\frac{1}{q}B^q\right)U^*=|AB|$$

for some unitary matrix U if and only if $A^p=B^q$.

We conclude the paper with the following remarks concerning our refined matrix Young inequality (9).

Remark 1. For the case $p=q=2$, inequality (9) degenerates to an equality. In fact, it has been observed in [8] that if $A,B,X\in M_n(\mathbb{C})$, then

$$\left\|A^*AX+XB^*B\right\|_2^2=\left\|A^*AX-AB^*B\right\|_2^2+4\left\|AXB^*\right\|_2^2. \tag{10}$$

Remark 2. If $a,b\geqslant 0$ and $p,q>1$ with $\frac{1}{p}+\frac{1}{q}=1$, then it is easy to see that

$$\left(\frac{a^p}{p}+\frac{b^q}{q}\right)^2\geqslant\left(\frac{a^p}{p}+\frac{b^q}{q}-ab\right)^2+a^2b^2, \tag{11}$$

$$\left(\frac{a^p}{p}+\frac{b^q}{q}+ab\right)^2\geqslant\left(\frac{a^p}{p}+\frac{b^q}{q}-ab\right)^2+4a^2b^2 \tag{12}$$

and

$$\frac{1}{s}\left|a^p-b^q\right|\geqslant\frac{a^p}{p}+\frac{b^q}{q}-ab, \tag{13}$$

where $s=\min\,(p,q)$.

Based on inequalities (11)–(13), one can argue as in the proof of inequality (9) to show that if $A,B,X\in M_n(\mathbb{C})$ such that A and B are positive semidefinite and if $p,q>1$ with $\frac{1}{p}+\frac{1}{q}=1$, then

$$\left\|\frac{1}{p}A^pX+\frac{1}{q}XB^q\right\|_2^2\geqslant\left\|\frac{1}{p}A^pX+\frac{1}{q}XB^q-AXB\right\|_2^2+\|AXB\|_2^2, \tag{14}$$

$$\left\|\frac{1}{p}A^pX+\frac{1}{q}XB^q+AXB\right\|_2^2 \geqslant \left\|\frac{1}{p}A^pX+\frac{1}{q}XB^q-AXB\right\|_2^2 + 4\ \|AXB\|_2^2, \tag{15}$$

$$\frac{1}{s}\left\|A^pX-XB^q\right\|_2 \geqslant \left\|\frac{1}{p}A^pX+\frac{1}{q}XB^q-AXB\right\|_2, \tag{16}$$

where $s=\min(p,q)$.

Note that inequality (14) is another refinement of inequality (5). When comparing between the refined inequalities (9) and (14), one should bear in mind that for $p=q=2$, inequality (9), which becomes an equality in this case, is better than inequality (14). However, for $p \neq q$, none of these inequalities is uniformly better than the other. It should be also noted that inequality (15) improves upon inequality (2.2) in [5] and the Hilbert–Schmidt norm version of the case $t=2$ of inequality (5) in [14], and that inequality (16) can be utilized to give direct proofs of the “if ” parts in Corollaries 1–3.

Remark 3. Inequality (9) can be generalized to pairs of commuting positive semidefinite matrices. Let $A, B, C, D, X \in M_n(\mathbb{C})$ such that A, B, C, D are positive semidefinite, $AD=DA$, and $BC=CB$. If $p, q>1$ with $\frac{1}{p}+\frac{1}{q}=1$, then

$$\left\|\frac{1}{p}A^pXC^p+\frac{1}{q}D^qXB^q\right\|_2^2 \geqslant \frac{1}{r^2}\left\|A^pXC^p-D^qXB^q\right\|_2^2 + \|ADXCB\|_2^2, \tag{17}$$

where $r=\max(p,q)$. In particular, if $C=D=I$, then we retain inequality (9). In view of the fact that commuting positive semidefinite matrices are simultaneously unitarily diagonalizable, inequality (17) can be derived by a slight modification of the proof of inequality (9).

Remark 4. Inequality (17) can be extended to arbitrary (i.e., not necessarily positive semidefinite) pairs of doubly commuting matrices by applying it to the absolute values of these matrices. Let $A, B, C, D, X \in M_n(\mathbb{C})$ such that $AD=DA$, $A^*D=DA^*$, $BC=CB$, and $B^*C=CB^*$. If $p, q>1$ with $\frac{1}{p}+\frac{1}{q}=1$, then

$$\left\|\frac{1}{p}|A|^p\,X\,|C|^p+\frac{1}{q}|D|^q\,X\,|B|^q\right\|_2^2 \geqslant \frac{1}{r^2}\left\||A|^p\,X\,|C|^p-|D|^q\,X\,|B|^q\right\|_2^2+\left\|ADXB^*C^*\right\|_2^2, \tag{18}$$

where $r = \max(p, q)$. In particular, if $C = D = I$, then we have

$$\left\| \frac{1}{p} |A|^p X + \frac{1}{q} X |B|^q \right\|_2^2 \geqslant \frac{1}{r^2} \left\| |A|^p X - X |B|^q \right\|_2^2 + \left\| AXB^* \right\|_2^2, \tag{19}$$

which is a natural extension of inequality (9) to arbitrary matrices.

To prove inequality (18), we note that, based on the spectral theorem, the commutation assumptions on A, B, C, D imply that $|A|\,|D| = |AD|$ and $|C|\ |B| = |CB|$. To see that $\||AD|\, X\, |CB|\|_2 = \|ADXB^*C^*\|_2$, we need to invoke (2) and the polar decompositions of AD and CB. This can also be seen by employing Lemma 6 in [9], which insures that

$$s_j\left(|AD|\, X\, |CB|\right) = s_j\left(ADXB^*C^*\right) \tag{20}$$

for $j = 1, 2, \ldots, n$.

We remark here that it is possible to give generalizations of inequalities (14)–(16) analogous to the generalizations (17)–(19) of inequality (9).

Remark 5. It is known that inequality (5) does not hold for the general class of unitarily invariant norms (see [1,6,12]). We infer from this that our generalized matrix Young inequality (18) also does not hold for this class of norms. In spite of the failure of inequality (5) for the general class of unitarily invariant norms, some weak matrix Young inequalities that are valid for this class of norms have been recently obtained in [12].

Remark 6. Although we have confined our discussion to matrices, considered as operators on a finite-dimensional Hilbert space, our generalized matrix Young inequality (18) can be easily extended to operators acting on an infinite-dimensional separable complex Hilbert space. In view of Remark 3, it is sufficient to extend inequality (17) to this setting. This can be achieved by appealing to a theorem of Voiculescu [13], which insures that if T_1 and T_2 are commuting positive semidefinite operators, then there are diagonal operators Λ_1 and Λ_2 and a unitary operator U such that both $T_1 - U\Lambda_1U^*$ and $T_2 - U\Lambda_2U^*$ are Hilbert–Schmidt operators with arbitrarily small Hilbert–Schmidt norm.

References

[1] T. Ando, Majorizations and inequalities in matrix theory, Linear Algebra Appl. 199 (1994) 17–67.
[2] T. Ando, Matrix Young inequalities, Oper. Theory Adv. Appl. 75 (1995) 33–38.
[3] R. Bhatia, Matrix Analysis, Springer, New York, 1997.
[4] R. Bhatia, F. Kittaneh, On singular values of a product of operators, SIAM J. Matrix Anal. Appl. 11 (1990) 271–277.
[5] R. Bhatia, F. Kittaneh, Notes on matrix arithmetic–geometric mean inequalities, Linear Algebra Appl. 308 (2000) 77–84.

[6] R. Bhatia, K.R. Parthasarathy, Positive definite functions and operator inequalities, Bull. London Math. Soc. (2000).
[7] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, New York, 1985.
[8] F. Kittaneh, A note on the arithmetic–geometric mean inequality for matrices, Linear Algebra Appl. 171 (1992) 1–8.
[9] F. Kittaneh, Norm inequalities for fractional powers of positive operators, Lett. Math. Phys. 27 (1993) 279–285.
[10] F. Kittaneh, On some operator inequalities, Linear Algebra Appl. 208/209 (1994) 19–28.
[11] F. Kittaneh, Some norm inequalities for operators, Canad. Math. Bull. 42 (1999) 87–96.
[12] H. Kosaki, Arithmetic-geometric mean and related inequalities for operators, J. Funct. Anal. 156 (1998) 429–451.
[13] D. Voiculescu, Some results on norm ideal perturbations of Hilbert space operators, J. Operator Theory 2 (1979) 3–37.
[14] X. Zhan, Inequalities for unitarily invariant norms, SIAM J. Matrix Anal. Appl. 20 (1999) 466–470.